Defense Threat Reduction Agency
8725 John J. Kingman Road, MS 6201
Fort Belvoir, VA 22060-6201

DTRA-TR-06-34

TECHNICAL REPORT

Grid-Search Techniques for Seismic Event Location and Phase Association

Approved for public release; distribution is unlimited.

March 2007

DTRA01-00-C-0102

Nafi Toksoz

Prepared by:
Massachusetts Institute of Technology
Department of Earth, Atmospheric and Planetary Sciences
Earth Resources Laboratory
Cambridge, MA 02139

DISTRIBUTION LIST UPDATE

This mailer is provided to enable DTRA to maintain current distribution lists for reports. (We would appreciate you providing the requested information.)

- ❑ Add the individual listed to your distribution list.
- ❑ Delete the cited organization/individual.
- ❑ Change of address.

Note:
Please return the mailing label from the document so that any additions, changes, corrections or deletions can be made easily. For distribution cancellation or more information call DTRA/CST (703) 767-4725.

NAME: _____

ORGANIZATION: _____

OLD ADDRESS	NEW ADDRESS
_____	_____
_____	_____
_____	_____

TELEPHONE NUMBER: ( ) _____

DTRA PUBLICATION NUMBER/TITLE	CHANGES/DELETIONS/ADDITONS, etc. *(Attach Sheet if more Space is Required)*
_____	_____
_____	_____
_____	_____

DTRA or other GOVERNMENT CONTRACT NUMBER: _____

CERTIFICATION of NEED-TO-KNOW BY GOVERNMENT SPONSOR (if other than DTRA):

SPONSORING ORGANIZATION: _____

CONTRACTING OFFICER or REPRESENTATIVE: _____

SIGNATURE: _____

DESTRUCTION NOTICE

FOR CLASSIFIED documents, follow the procedures in DoD 5550.22-M, National Industrial Security Program Operating Manual, Chapter 5, Section 7 (NISPOM) or DoD 5200.1-R, Information Security Program Regulation, Chapter 1X.

FOR UNCLASSIFIED limited documents, destroyed by any method that will prevent disclosure of contents or reconstruction of the document.

Retention of this document by DoD contractors is authorized in accordance with DoD 5220.22-M, Industrial Security Manual.

PLEASE NOTIFY THE DEFENSE THREAT REDUCTION AGENCY, ATTN: CST, 8725 JOHN J. KINGMAN ROAD, STOP-6201, FT BELVOIR, VA 22060-6201, IF YOUR ADDRESS IS INCORRECT, IF YOU WISH IT DELETED FROM THE DISTRIBUTION LIST, OR IF THE ADDRESSEE IS NO LONGER EMPLOYED BY YOUR ORGANIZATION.

DEFENSE THREAT REDUCTION AGENCY
ATTN: CST
8725 John J Kingman Road, MS 6201
Fort Belvoir, VA 22060-6201

DEFENSE THREAT REDUCTION AGENCY
ATTN:CST
8725 John J Kingman Road, MS 6201
Fort Belvoir, VA 22060-6201

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. **PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.**

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
00-01-2007	Technical report	

4. TITLE AND SUBTITLE
Grid-Search Techniques for Seismic Event Location and Phase Association

5a. CONTRACT NUMBER
DTRA01-00-C-0102

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER
463D

6. AUTHOR(S)
Nafi Toksoz, William Rodi and Sudipta Sarkar

5d. PROJECT NUMBER
CD

5e. TASK NUMBER
CD, 2U

5f. WORK UNIT NUMBER
DH01869

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) AND ADDRESS(ES)
Massachusetts Institute of Technology
Department of Earth, Atmospheric and Planetary Sciences
Earth Resources Laboratory
Cambridge, MA 02139

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Defense Threat Reduction Agency
8725 John J. Kingman Road
MS 6201
Fort Belvoir, VA 22060-6201
CW/S. Mangino

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
DTRA-TR-06-34

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES
This work was sponsored by the Defense Threat Reduction Agency under RDT&E RMC code B 463D D K540 CD CD 01869 25904D.

14. ABSTRACT
The objectives of this project are to develop improved methods for locating seismic events and to develop new approaches to analyzing the seismic event location uncertainty. Our efforts are focused on the problems of both single and multiple event location. For a single event, we have formulated a statistical approach based on maximum-likelihood estimation and implemented it with grid-search and Monte Carlo techniques to yield an algorithm called "Grid-Search Single Event Location" (GSEL), which computes non-elliptical confidence regions on event locations for a wide class of assumed probability distributions (Gaussian and non-Gaussian) and for data picking errors. We have applied GSEL to regional arrival data from the 1991 Racha earthquake sequence and from the 1998 Adana (Turkey) earthquake sequence. Our multiple-event grid-search location algorithm (called GMEL) is also based on a maximum likelihood formulation of the location problem and it solves jointly for the location parameters (hypocenters and origin times) of seismic events in a cluster and travel-time corrections at the stations recording the events, GMEL accommodates non-Gaussian as well as Gaussian models of picking errors. We test this GMEL with clusters of 1999 Izmit/Duzce (Turkey) earthquake sequence. The approach we developed is based on the maximum-likelihood framework determines confidence regions on event locations under a general class of picking error models (Gaussian and non-Gaussian).

15. SUBJECT TERMS
Seismic events Gaussian Grid-Search Single Event Location GMEL

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			19b. TELEPHONE NUMBER *(include area code)*
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	45	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39.18

CONVERSION TABLE

Conversion Factors for U.S. Customary to metric (SI) units of measurement.

MULTIPLY →	BY →	TO GET
TO GET ←	BY ←	DIVIDE

angstrom	1.000 000 x E -10	meters (m)
atmosphere (normal)	1.013 25 x E +2	kilo pascal (kPa)
bar	1.000 000 x E +2	kilo pascal (kPa)
barn	1.000 000 x E -28	meter^2 (m^2)
British thermal unit (thermochemical)	1.054 350 x E +3	joule (J)
calorie (thermochemical)	4.184 000	joule (J)
cal (thermochemical/cm^2)	4.184 000 x E -2	mega joule/m^2 (MJ/m^2)
curie	3.700 000 x E +1	*giga bacquerel (GBq)
degree (angle)	1.745 329 x E -2	radian (rad)
degree Fahrenheit	$t_k = (t°f + 459.67)/1.8$	degree kelvin (K)
electron volt	1.602 19 x E -19	joule (J)
erg	1.000 000 x E -7	joule (J)
erg/second	1.000 000 x E -7	watt (W)
foot	3.048 000 x E -1	meter (m)
foot-pound-force	1.355 818	joule (J)
gallon (U.S. liquid)	3.785 412 x E -3	meter^3 (m^3)
inch	2.540 000 x E -2	meter (m)
jerk	1.000 000 x E +9	joule (J)
joule/kilogram (J/kg) radiation dose absorbed	1.000 000	Gray (Gy)
kilotons	4.183	terajoules
kip (1000 lbf)	4.448 222 x E +3	newton (N)
kip/inch^2 (ksi)	6.894 757 x E +3	kilo pascal (kPa)
ktap	1.000 000 x E +2	newton-second/m^2 ($N-s/m^2$)
micron	1.000 000 x E -6	meter (m)
mil	2.540 000 x E -5	meter (m)
mile (international)	1.609 344 x E +3	meter (m)
ounce	2.834 952 x E -2	kilogram (kg)
pound-force (lbs avoirdupois)	4.448 222	newton (N)
pound-force inch	1.129 848 x E -1	newton-meter (N-m)
pound-force/inch	1.751 268 x E +2	newton/meter (N/m)
pound-force/foot^2	4.788 026 x E -2	kilo pascal (kPa)
pound-force/inch^2 (psi)	6.894 757	kilo pascal (kPa)
pound-mass (lbm avoirdupois)	4.535 924 x E -1	kilogram (kg)
pound-mass-foot^2 (moment of inertia)	4.214 011 x E -2	kilogram-meter^2 ($kg-m^2$)
pound-mass/foot^3	1.601 846 x E +1	kilogram-meter^3 (kg/m^3)
rad (radiation dose absorbed)	1.000 000 x E -2	**Gray (Gy)
roentgen	2.579 760 x E -4	coulomb/kilogram (C/kg)
shake	1.000 000 x E -8	second (s)
slug	1.459 390 x E +1	kilogram (kg)
torr (mm Hg, 0^0 C)	1.333 22 x E -1	kilo pascal (kPa)

*The bacquerel (Bq) is the SI unit of radioactivity; 1 Bq = 1 event/s.
**The Gray (GY) is the SI unit of absorbed radiation.

TABLE OF CONTENTS

1. SUMMARY

The objectives of this project are to develop improved methods for locating seismic events and to develop new approaches to analyzing the seismic event location uncertainty. Our efforts are focused on the problems of both single and multiple event location.

For a single event, we have formulated a statistical approach based on maximum-likelihood estimation and implemented it with grid-search and Monte Carlo techniques to yield an algorithm called "Grid-Search Single Event Location" (GSEL), which computes non-elliptical confidence regions on event locations for a wide class of assumed probability distributions (Gaussian and non-Gaussian), and for data picking errors. We have applied GSEL to regional arrival data from the 1991 Racha earthquake sequence and from the 1998 Adana (Turkey) earthquake sequence.

Our multiple-event grid search location algorithm (called GMEL) is also based on a maximum-likelihood formulation of the location problem and it solves jointly for the location parameters (hypocenters and origin times) of seismic events in a cluster and travel-time corrections at the stations recording the events. GMEL accommodates non-Gaussian as well as Gaussian models of picking errors. We test this GMEL with clusters of 1999 Izmit/Duzce (Turkey) earthquake sequence.

The approach we developed is based on the maximum-likelihood framework determines confidence regions on event locations under a general class of picking error models (Gaussian and non-Gaussian).

2. INTRODUCTION

The objective of this project is to develop new techniques for seismic event location that yield improved location estimates and reliable estimates of location uncertainty for small, sparsely recorded events. We apply grid-search and other numerical techniques with the specific objectives of (1) obtaining globally optimal event locations from sparse data under a general error model; (2) performing a rigorous uncertainty analysis on seismic event locations; and (3) increasing the number of seismic phases which constrain event locations by integrating phase association into the event location algorithm.

The first phase of the project focused on the problem of characterizing the uncertainty in seismic event locations. Conventional methods for inferring elliptically-shaped confidence regions on event locations (e.g., Flinn, 1965) employ assumptions that may not be valid in the case of small, sparsely recorded events. These assumptions include local linearity of the travel-time vs. hypocenter forward model and the treatment of errors in the arrival time data as uncorrelated, Gaussian random variables. The linear approximation is only appropriate when the "true" confidence region for an event is small compared to the distances to stations, and when it does not encompass large velocity variations in the Earth. Gaussian error models do not capture the observational difficulties of correctly identifying and picking low signal-to-noise arrivals (Jeffreys, 1932), and are ad hoc representations of what is often a larger source of error: the errors in the travel-time tables used in the location process. We address these difficulties by developing a general theoretical approach to event location uncertainty in terms of likelihood functions and hypothesis testing, and implemented the approach with grid-search and Monte Carlo techniques (Rodi and Toksöz, 2000, 2001). The resulting algorithm computes confidence regions on event locations that account for such complexities as nonlinearity of the forward travel-time problem and non-Gaussian distributions of observational (picking) errors.

A major objective of our project was to also incorporate a rigorous treatment of the errors in travel-time models (modeling errors) into our uncertainty analysis. Toward

this goal, we have linked our uncertainty analysis to the process of seismic travel-time calibration, recognizing that modeling errors are due to the uncertainty in calibration parameters such as path corrections and tomographic velocity models. Therefore, we have extended our formulation to the multiple-event location problem, i.e. whereby data from multiple events and stations are used to simultaneously infer the locations of the events and calibration parameters. The tradeoffs and correlations among the unknowns of this joint inverse problem effect an implicit account of modeling errors. Our first step in implementing this formulation was to extend our single-event grid-search location algorithm (GSEL) to a multiple-event one (GMEL). This algorithm is described in this report and is also compared to some other multiple-event location algorithms stemming back to the early work of Douglas (1967), Lilwall and Douglas (1970), and Dewey (1971). Finally, we describe our formulation of multiple-event location uncertainty and our efforts to implement in an extended version of GMEL.

3. SINGLE EVENT LOCATION

Our approach to seismic event location is based on a maximum-likelihood formulation. We define an optimal location for a seismic event to be that which maximizes a likelihood function, constructed on the basis of an assumed statistical model of errors in the seismic data used in locating the event. Confidence regions are defined in terms of hypothesis tests using likelihood ratios as the test statistics. We have formulated and implemented this approach for three types of seismic data used in nuclear monitoring: arrival times, azimuths and slownesses. To simplify the discussion here, we consider only arrival times.

3.1. Single-Event Location Formulation

The hypocentral parameters of a seismic event are an origin time t and location $\mathbf{x} \equiv \theta\, \phi, z)$, where θ is latitude, ϕ is longitude and z is depth. Let $\mathbf{d} = (d_1, d_2, \ldots, d_n)$ be an n-dimensional vector of arrival times picked from various seismic phases at a seismic network. The event location problem may be expressed as

$$d_i = t + T_i(\mathbf{x}) + c_i + e_i, \qquad i = 1,\ldots..,n. \tag{1}$$

T_i is a travel-time function for the ith datum, which in our algorithms is evaluated by interpolating a travel-time table that samples travel-time as a function of epicentral distance and event depth (for a 1-D Earth model) or as a function of **x** (for a 3-D Earth model). The term e_i denotes the observational, or picking, error in d_i. The term c_i can be interpreted in two ways. First, it is a correction to $T_i(\mathbf{x})$, accounting for the difference between the travel-time function and the true travel-times in the Earth. However, if c_i is not known, it can be considered an additional source of error in d_i, and from this view it is sometimes referred to as a "modeling error." In our formulation, we assume that any known corrections have been included in T_i, leaving c_i as an unknown error. The uncertainty analysis for the event hypocenter **x** must account for the combined error, $c_i + e_i$.

Our current algorithms assume the picking errors are statistically independent and, following Billings et al (1994), that each is distributed with an exponential power distribution, whose probability density function (p.d.f.) is given by

$$f(e_i) = \frac{1}{K(p)\sigma_i} \exp\left\{ -\frac{1}{p} \left| \frac{e_i}{\sigma_i} \right|^p \right\} \tag{2}$$

where the scale parameter σ_i is a standard error, $K(p)=2p^{1/p}\,\Gamma(1+1/p)$, and Γ is the gamma function. When $p = 2$, then e_i is normally distributed with a mean of zero and variance of $(\sigma_i)^2$. When $p = 1$, it is exponentially distributed. We assume that the standard errors are known in a relative sense and write

$$\sigma_i = \sigma \nu_i \tag{3}$$

where the ν_i are known but the universal scale parameter, σ is not.

We currently assume that the modeling errors, c_i, are deterministic unknowns. The full set of unknowns in our formulation of the single-event location problem thus comprise **x**, t, σ and $\mathbf{c} = (c_1, c_2\ldots,c_n)$.

3.2. Maximum Likelihood Formulation

The joint p.d.f. of the n data is the product of the error p.d.f.'s. Considered as a function of the unknown parameters ($\mathbf{x}$, t, and σ) this joint p.d.f. is the likelihood function we seek to maximize. Denoting it as $L(\mathbf{x}, t, \sigma, \mathbf{c}; \mathbf{d})$, our assumptions imply

$$-\log L(\mathbf{x},t,\sigma,\mathbf{c};\mathbf{d}) = \sum_{i=1}^{n} \log \nu_i + n \log K(p) + n \log \sigma + \frac{1}{p\sigma^p} + \frac{1}{p\sigma^p} \Psi(\mathbf{x},t,\mathbf{c};\mathbf{d}) \qquad (4)$$

where the "data misfit" function Ψ is given by

$$\Psi(\mathbf{x},t,\mathbf{c};\mathbf{d}) = \sum_{i=1}^{n} \left| d_i - t - T_i(\mathbf{x}) - c_i \right|^p / (\nu_i)^p \qquad (5)$$

We show the maximum-likelihood estimates of the unknowns as $\mathbf{x}_{\text{ml}}$, t_{ml}, σ_{ml} and $\mathbf{c}_{\text{ml}}$. The maximization may be subjected to prior constraints on the parameters. For the applications here the constraints of interest are

$$0 \le z \le z^{\max} \qquad (6)$$

$$\sigma^{\min} \le \sigma \le \sigma^{\max} \qquad (7)$$

$$c_i^{\min} \le c_i \le c_i^{\max} \qquad (8)$$

From the point of view of finding $\mathbf{x}$, t and $\mathbf{c}$, maximizing L is equivalent to minimizing Ψ, which is just a weighted p-norm of data residuals (to the power p). For example, with $p = 2$ the problem reduces to nonlinear least squares. However, with error models more general than we consider here, allowing for example asymmetric or multi-modal error distributions, the likelihood function would not necessarily be in terms of a simple residual norm.

Given its structure, L is amenable to a hierarchical maximization with respect to the unknown parameters. We define a "reduced" likelihood function which, for each fixed hypocenter, is maximum with respect to t, σ and $\mathbf{c}$ (subject to prior constraints):

$$\tilde{L}(\mathbf{x};\mathbf{d}) = \max_{t,\sigma,\mathbf{c}} L(\mathbf{x},t,\sigma,\mathbf{c};\mathbf{d}). \qquad (9)$$

For general $p \ge 1$, the maximization over t, σ and $\mathbf{c}$ can be performed with a combination of analytical and root-finding techniques. The location problem reduces to maximization of $\tilde{L}$ with respect to $\mathbf{x}$.

3.3. Grid Search

Our grid search algorithm computes the reduced likelihood function, $\tilde{L}$ in equation (9), at each point in a 3-D grid of hypocenters. Following previous workers, the hypocenter grid is constructed dynamically through a process of successive refinement. Our procedure for grid refinement resembles that of the 'neighborhood' search algorithm developed by Sambridge (1999). If the search is not restricted to a specified region, the first grid covers the entire globe and 0 to 700 km in depth at a coarse spacing: 100 km in depth, 9° in latitude, and 9° in longitude near the equator and increasing at higher latitudes. On each pass of grid refinement, nodes are added as neighbors of a subset of grid points comprising the "best" (largest $\tilde{L}$) points tested thus far. Neighbors are placed at one-third the grid-spacing of the previous pass. The size of the grid subset chosen for refinement is reduced on each pass. The search ends when the grid spacing is less than 0.3 km.

3.4. Non-Elliptical Confidence Regions

In conventional event location algorithms, confidence regions on the hypocenter, epicenter and depth of an event are computed under the assumptions that the data errors are Gaussian $p = 2$ and that the travel-time functions, $T_i(\mathbf{x})$, are well-approximated as locally linear near $\mathbf{x} = \mathbf{x}_{ml}$. These assumptions lead to hypocenter and epicenter confidence regions that are elliptical in shape. The size of the confidence regions, for a given confidence level, is scaled by a critical value of the F distribution for an appropriate number of degrees of freedom, as determined by a prior assumption about the data variance, σ^2. This approach is developed by Flinn (1965), Evernden (1969) and Jordan and Sverdrup (1981) under differing assumptions about σ^2. (unknown, known and partially known, respectively).

We have generalized this approach to avoid the linearity approximation and to allow for non-Gaussian data errors and arbitrary parameter constraints. As in the conventional approach, we define a confidence region in terms of a test statistic, τ, which is a function of the data and parameters being tested. For a hypocenter confidence

region, we write the statistic as τ(d, **x**). The confidence region, at confidence level β (e.g., β = 90 %), comprises those values of **x** that satisfy the inequality

$$F[\tau(\mathbf{d},\mathbf{x})] \le \beta \tag{10}$$

where **d** is the observed data vector, and $F[\ ]$ denotes the cumulative distribution function (c.d.f.) of a random variable. Following a well-known statistical approach, we define the test statistic as the logarithm of a likelihood ratio:

$$\tau(\mathbf{d},\mathbf{x}) = \log \tilde{L}(\mathbf{x}_{ml};\mathbf{d}) - \log \tilde{L}(\mathbf{x};\mathbf{d}) = \log\left[\frac{\max_{\mathbf{x},t,\sigma,\mathbf{c}} L(\mathbf{x},t,\sigma,\mathbf{c};\mathbf{d})}{\max_{t,\sigma,\mathbf{c}} L(\mathbf{x},t,\sigma,\mathbf{c};\mathbf{d})}\right] \tag{11}$$

That is, for given **x**, τ compares the difference between the maximum likelihood (achieved by $\mathbf{x}_{ml}$) and the likelihood achieved by **x**. Since the inequality (10) rejects large values of τ, confidence regions will exclude hypocenters that achieve relatively small values of likelihood, i.e., poor fits to the data. We point out that under the Gaussian, linear assumption, the likelihood ratio statistic is equivalent to the variance ratio on which elliptical confidence regions are based.

The statistic for a 2-D confidence region on the event epicenter, (θ, ϕ), is defined by

$$\tau(\mathbf{d},\theta,\phi) = \log \tilde{L}(\mathbf{x}_{ml};\mathbf{d}) - \log \max_{z} \tilde{L}(\theta,\phi,z;\mathbf{d}) \tag{12}$$

and for a confidence interval on focal depth we use

$$\tau(\mathbf{d},z) = \log \tilde{L}(\mathbf{x}_{ml};\mathbf{d}) - \log \max_{\theta,\phi} \tilde{L}(\theta,\phi,z;\mathbf{d}) \tag{13}$$

Confidence regions using these likelihood-ratio statistics could still be defined via the inequality of equation (10), except this inequality presumes that the distribution (c.d.f.) of τ does not depend on the true values of the parameters that are not being tested. To address this, we assume the main dependence is on focal depth and σ, and rewrite the c.d.f. of τ as $F[\tau, z, \sigma]$. We generalize the inequality of equation (10) to use the c.d.f. of τ that is minimum with respect to the untested parameters. Thus, the hypocentral confidence region is given by

$$\min_{\sigma} F[\tau(\mathbf{d},\mathbf{x}),z,\sigma] \le \beta \tag{14}$$

The epicentral confidence region is

$$\min_{\sigma,z} F[\tau(\mathbf{d},\theta,\phi); z,\sigma] \le \beta \tag{15}$$

and the focal depth confidence interval is

$$\min_{\sigma} F[\tau(\mathbf{d},\theta,\phi); z,\sigma] \le \beta \tag{16}$$

With these definitions, a confidence region will include the true value of the tested parameters at least 100β percent of the time.

3.5. Monte Carlo Confidence Levels

Without assumptions like linearity of T_i, it is not possible to derive an analytic expression for the c.d.f. of the test statistic τ, which is needed to compute confidence regions. However, we can approximate the c.d.f. using Monte Carlo simulation. We outline the technique for hypocentral confidence regions. The idea is to compute τ for many randomly generated samples of the error vector. We generate each error, e_i^{mc}, using a pseudo-random number generator in accordance with the assumed error distribution (eqs. (2) – (3)) for some given "true" σ. Then, for given true hypocentral parameters, **x** and t, and station corrections, **c**, synthetic data are calculated as

$$d_i^{mc} = t + T_i(\mathbf{x}) + c_i + e_i^{mc} \tag{17}$$

We apply our grid search algorithm to these data to obtain the maximum-likelihood hypocenter, $\mathbf{x}_{ml}^{mc}$. Plugging this into the formula for τ,

$$\tau(\mathbf{d}^{mc},\mathbf{x}) = \log \tilde{L}(\mathbf{x}_{ml}^{mc};\mathbf{d}^{mc}) - \log \tilde{L}(\mathbf{x};\mathbf{d}^{mc}) \tag{18}$$

we obtain one sample from $F[\tau; z, \sigma.]$. We compare this sample to the observed value of the statistic, $\tau(\mathbf{d}, \mathbf{x})$, obtained from the real data. We count a rejection of **x** if

$$\tau(\mathbf{d}^{mc},\mathbf{x}) < \tau(\mathbf{d},\mathbf{x}) \tag{19}$$

The proportion of rejections after many Monte Carlo trials yields an estimate of $F[\tau(\mathbf{d},\mathbf{x}); z, \sigma.]$. Performing this simulation for multiple values of $\sigma.$ and then minimizing amongst them gives the *lowest* confidence level such that the confidence region includes **x**.

The process for depth confidence intervals and epicenter confidence regions proceeds in the same manner, except that in the latter case the simulation is performed for

multiple values of true depth as well as $\sigma_.$, and the confidence level is minimized over both.

3.6. Examples of Confidence Regions

Racha earthquake sequence

We show some examples of non-elliptical confidence regions computed with our Monte Carlo technique. Figure 1 shows confidence regions computed for two events from the 1991 Racha, Georgia, earthquake sequence, studied by Myers and Schultz (2000). The data sets comprise P wave arrival times from 5 or 6 regional stations, obtained from the International Seismological Centre (ISC) but with two stations repicked by Lawrence Livermore National Laboratory (LLNL). The confidence regions were computed under the Gaussian error model ($p = 2$) with data standard deviations fixed to 1.0 sec for the ISC picks and 0.5 sec for LLNL picks. We used the IASP91 travel-time tables, and fixed the travel-time corrections to zero ($c_i = 0$).

The results illustrate the effect of nonlinearity on the confidence regions of sparsely recorded events. The confidence regions depart significantly from an elliptical shape. This is due largely to the fact that the depth of these events is poorly constrained by first arrivals from a few (5 or 6) stations covering a limited distance range (6° to 22°). Travel-time does not behave as a linear function over the wide range of event depths that is consistent with the data.

We see also from Figure 1 that the confidence regions do not include the local-network locations, except at very high confidence levels. This was generally the case for 18 *Racha* events we analyzed, all of which had six or fewer data. For the 14 of these events whose epicenters were reasonably constrained, the locations were consistently mislocated by roughly 40 km north-northwest of the respective ground-truth locations (see Figure 2).

These results are consistent with those of Myers and Schultz (2000), who showed that the mislocations are due to the need for travel-time corrections as large as 3 seconds at some stations.

Adana (Turkey) earthquake sequence

Figure 3 shows confidence regions determined for the 27 June 1998 M=6.2 *Adana* earthquake and its largest aftershock of July 4, 1998. The data used in these examples are P and Pn arrival times reported in the Reviewed Event Bulletin (REB) of the International Data Centre (IDC); secondary arrivals were omitted for this study. The stations ranged from regional to far-teleseismic distances and numbered 24 for the mainshock and 20 for the aftershock. Gaussian errors were assumed with a standard error (σ_i) of 1 sec.

In contrast to the *Racha* examples, confidence regions are more elliptical in shape. This is because the confidence regions are small compared to the distances to the stations, and they are contained mainly in the upper mantle where there are no large velocity contrasts (in the IASP91 model). Like the *Racha* examples, however, we see that only the confidence regions at high confidence levels cover the true locations of the *Adana* events, as inferred from a dense local network that surrounded the events (Aktar et al., 2000). The fact that the mislocation is similar (west and deep) for the two events suggests it is not due to picking errors that were anomalously large compared to our assumed picking accuracy.

3.7. Approach to Modeling Errors

In the examples above, reasonable assumptions about picking errors do not yield valid conclusions about the uncertainty in event locations. The confidence regions do not include the true event locations except at very high confidence levels, and the mislocation is similar for different events in the same region, suggesting a repeatable source of error. The reason is that some stations need large travel-time corrections (c_i). When these corrections are not known, our uncertainty analysis must account for them as "modeling errors."

A simple, and commonly used, way to do this is to inflate the assumed variance of the data errors, attempting thus to define the probability distribution of the sum of picking and modeling errors ($c_i + e_i$ in equation (1)). This would be appropriate if modeling errors were independent between stations and seismic phases. Another mechanism is to modify the shape, as well as the width, of the error distribution, e.g., manipulating p in our

formulation. However, to the extent that modeling errors are correlated between stations and phases, the confidence regions that result might still not be indicative of the true location error. Ultimately, a statistical analysis of actual picking and modeling errors is needed to derive an appropriate error model.

Figures 4 and 5 show some of these ad hoc treatments of modeling errors for the two Adana events, i.e. (1) using an exponential distribution ($p = 1$) for the picking errors (instead of Gaussian); (2) inflating the variance of picking errors ($\sigma_i = 2$ sec instead of 1 sec); and (3) allowing (partially) unknown travel-time corrections at the stations ($-2 \text{ sec} \leq c_i \leq 2$ sec instead of $c_i = 0$). We see that all three alternative error models lead to larger confidence regions, although only slightly so in the first case (exponential distribution). For the mainshock, the second and third approaches include the ground-truth location at reasonable confidence levels, but not for the aftershock. In general, we cannot expect these approaches to provide consistent, valid uncertainty estimates.

4. MULTIPLE EVENT LOCATION

Multiple event location methods solve jointly for the location parameters of events in a cluster and station corrections. This method fits travel-time residuals observed from many events with parameterized corrections. An analysis of the errors in the correction estimates provides the statistical model of modeling errors that we need. In the following sections we will discuss our grid-search multiple event location method (GMEL) and compare it with other similar approaches.

4.1. Multiple-Event Location Formulation

GMEL processes arbitrary data sets of seismic arrival times, azimuths and slownesses observed from multiple events for multiple stations and phases. To simplify the discussion here, we consider data sets comprising only arrival times.

We consider arrival times from m seismic events and n seismic stations. We denote the origin time and hypocenter of the jth event as t_j and $\mathbf{x}_j$, respectively. As a further simplification for this write-up, we assume there is at most one arrival time

observed for each event-station pair, and denote this time as d_{ij} for the ith station and jth event. It is understood that d_{ij} exists only for a subset of the mn possibilities. The multiple-event location problem can then be written as

$$d_{ij} = t_j + T_i(\mathbf{x}_j) + c_{ij} + e_{ij} \,. \tag{20}$$

Here, T_i is a known travel-time function for the ith station, obtained from an assumed Earth model; c_{ij} is an unknown correction to this function that accounts for differences between the real Earth and the assumed model; and e_{ij} is an observational ("picking") error.

The multiple-event location problem is to solve equation (20) for the location parameters ($\mathbf{x}_j$, t_j) of the m events jointly with the path corrections (c_{ij}). However, with no constraints on the path corrections, this problem is hopelessly ill posed. Therefore, GMEL and many previous methods assume that the path corrections at a given station are the same for all events, which is tantamount to assuming that the events are in a small cluster. In this case, we can replace c_{ij} with c_i and re-write (20) as

$$d_{ij} = t_j + T_i(\mathbf{x}_j) + c_i + e_{ij} \,. \tag{21}$$

The problem unknowns are now the $\mathbf{x}_j$, t_j and c_i. The c_i are sometimes referred to as station corrections instead of path corrections, but it is important to realize they are dependent on the location of the event cluster.

4.2. GMEL

GMEL addresses the problem stated in equation (21) (with event-independent travel-time corrections) and solves jointly for the corrections and event location parameters. The error assumptions used are identical to those in GSEL (see Rodi and Toksöz, 2001), whereby each e_{ij} has a generalized Gaussian distribution of order p ($p = 2$ for Gaussian). In GMEL, the data standard errors, σ_{ij},, are assumed known in a relative sense, such that

$$\sigma_{ij} = \sigma_i \nu_{ij} \tag{22}$$

where the ν_{ij} are known, but the station-dependent scale parameter, σ_i are not. The error model implies a likelihood function, L, given by

$$-\log L = \text{const} + \sum_i n_i \log \sigma_i + \Psi(\mathbf{x}_1, t_1, \mathbf{x}_2, t_2, \ldots, c_1, c_2, \ldots) \tag{23}$$

where n_i is the number of events recorded at the ith station, and Ψ is a data misfit function. We can write Ψ as either a sum of station-specific or event-specific misfit functions, i.e.

$$\Psi = \sum_j \Psi_j^{\text{ev}}(\mathbf{x}_j, t_j, c_1, c_2, \ldots) = \sum_i \frac{1}{(\sigma_i)^p} \Psi_i^{\text{st}}(\mathbf{x}_1, t_1, \mathbf{x}_2 t_2, \ldots, c_i) \tag{24}$$

with

$$\Psi_j^{\text{ev}}(\mathbf{x}_j, t_j, c_1, c_2, \ldots) = \sum_i \frac{1}{(\sigma_i)^p} | d_{ij} - t_j - T_i(\mathbf{x}_j) - c_i |^p / (\nu_{ij})^p \tag{25}$$

$$\Psi_i^{\text{st}}(\mathbf{x}_1, t_1, \mathbf{x}_1, t_1, \ldots, c_i) = \sum_j | d_{ij} - t_j - T_i(\mathbf{x}_j) - c_i |^p / (\nu_{ij})^p . \tag{26}$$

GMEL maximizes the likelihood with respect to the event parameters ($\mathbf{x}_j$, t_j) and station parameters (c_i and σ_i), subject to a prior upper and lower bound on each parameter. The constraints on $\mathbf{x}_j$ comprise bounds on depth and a maximum epicentral distance from some specified geographic point.

We point out that with the generalized Gaussian error model, maximizing likelihood with respect to the event locations and station corrections, with the σ_i fixed, is equivalent to minimizing the data misfit, Ψ, and in the Gaussian case ($p = 2$) this is a nonlinear least-squares problem.

The scheme GMEL currently uses for maximizing likelihood is an iteration of a two-step process. First, fixing the station corrections and standard errors to their initial values (the average of their upper and lower bounds), each event is located in turn. This is performed by applying the single-event grid-search algorithm (GSEL) separately to each event. The second step is a loop over stations to estimate the station parameters (c_i and σ_i) with the event locations fixed. As in the event loop, the stations can be treated individually. In the Gaussian case with unrestrictive prior bounds, each station correction is simply the weighted mean residual at the station, but in general GMEL uses a root-finding procedure to find the correction that minimizes Ψ_i^{st}, the data misfit function for

the station. The event loops and station loops are repeated until the likelihood function stops achieving significant gains. We can summarize the GMEL algorithm as an iteration of the following:

1. For each j: minimize Ψ_j^{ev} with respect to $\mathbf{x}_j$ and t_j (with the c_i and $\sigma_{i\,i}$ fixed)
2. For each i: minimize Ψ_i^{st} with respect to c_i (with the $\mathbf{x}_j$ and t_j fixed)
3. For each i: maximize L with respect to $\sigma_{\cdot}$;

$$\sigma_{\cdot} = (\Psi_i^{st}/n_i)^{1/p}.$$

4.3. Theoretical Comparison of Methods

The multiple-event location methods we compare here are:

- JHD: joint hypocenter determination, version JHD89 (Dewey, 1971).
- HDC: hypocentral decomposition (Engdahl and Bergman, 2000; Jordan and Sverdrup, 1981).
- PMEL: progressive multiple-event location (Pavlis and Booker, 1983).
- DD: double-differencing (Waldhauser and Ellsworth, 2000).
- GMEL: grid-search multiple-event location.

We briefly describe the other methods in relation to GMEL. First, we note that PMEL, HDC and JHD all perform a joint least-squares inversion for event location parameters and event-independent station corrections; i.e. they minimize Ψ, defined in equations (24) through (26), with $p = 2$. DD does the same in a special case, which we will explain below. Second, we note that all of the methods except GMEL employ some sort of iterated linearized inversion algorithm for minimizing Ψ, with respect to location parameters, as opposed to grid-search in the case of GMEL. The essential differences between the algorithms have to do with the order in which the unknown parameters are solved for. This potentially makes a significant difference between algorithms because of the trade-offs among the unknown parameters in the multiple-event location problem. Jordan and Sverdrup (1981) showed that, under some simplifying approximations, the inverse problem (Equation 21) does not have a unique solution in that there is a perfect trade-off between a four-dimensional projection of the station corrections and the mean

hypocenter and origin time of the events. However, the relative locations amongst events and the complementary projection of the corrections can be determined uniquely, disallowing the effects of the data errors. However, relaxing their assumptions and allowing for data errors as well as possible deficiencies in the completeness of the data, the trade-offs between station corrections and event locations might be more complex.

We explained that GMEL minimizes the data misfit function with respect to the location of each event and the correction at each station in alternating loops over events and stations. JHD solves for all the parameters of all events and stations simultaneously in a linear inverse problem in which the travel-time functions, $T_i(\mathbf{x}_j)$, have been linearized about the current event locations. (The dependence on the t_j and c_i is exactly linear.) The solution of the linearized problem yields updates to event and station parameters that are added to their current values. The process is iterated, thus solving the nonlinear least-squares problem.

PMEL resembles GMEL in iterating over a two-step process, with one step updating the event locations and the other updating the station corrections. The events are relocated individually as in GMEL, but PMEL uses a conventional single-event location algorithm instead of grid search. The updating of station corrections is quite different from GMEL. PMEL solves for the c_i simultaneously in a linear system which has been pre-conditioned with an "annulling" or "projection" operator (Pavlis and Booker, 1980) that prevents the station corrections from fitting components of the travel-time residuals that can be fit with hypocenter changes on subsequent iterations. Thus, PMEL resolves trade-offs between station and event parameters in favor of the events. Since GMEL updates event locations prior to station corrections on each pass of its iteration, it also gives some preference to fitting the data with event parameters, but not as faithfully as PMEL.

HDC solves separately for the centroid location of the event cluster and the relative locations of the individual events with respect to this centroid ("cluster vectors"). It too repeats a two-step process, but again the steps are different from GMEL. The first step solves simultaneously for all the cluster vectors (relative locations) in the linear inverse problem resulting from linearizing the travel-time functions about the current absolute event locations. HDC pre-conditions this linear inverse problem in the opposite

sense of PMEL; i.e., it projects out the sensitivity to station corrections, which couples the problem across events (instead of stations as in PMEL) and which gives preference to the station corrections in fitting the travel-time residuals. (Under the Jordan-Sverdrup assumptions, this preference amounts to the constraint that the cluster vectors sum to zero.) In the second step of the HDC process, the updated cluster vectors are added to the current cluster centroid to get updated absolute locations, and then updated data residuals are computed. The station averages of these residuals are used to update the cluster centroid, and any remaining average residual at each station is its station correction. These two steps are repeated in an iteration.

Of the algorithms discussed thus far, DD follows the HDC paradigm most closely, but the step of finding relative locations is quite different. It solves for the relative locations in a linearized inverse problem obtained by forming arrival time differences, at each station, between pairs of close-by events recorded by the station. Wolfe (2002) has shown that in a simple case (uniform data weighting and all possible event pairs used in the inversion) this differencing procedure is equivalent to the HDC projection technique but with an implicit station-dependent re-weighting of the data (i.e., $\sigma_i^2 = 1/n_i$). In addition to omitting time differences for certain event pairs, another distinguishing feature of DD is that it can weight each difference in accordance with the distance between the two events. As a result of these features (omitting and distance-weighting of time differences), DD relaxes the assumption that path corrections (c_{ij}) are event-independent and, accordingly, does not precisely minimize the data misfit function, Ψ, defined above in equations (24) through (26).

The various algorithms compared here also differ with regard to two other issues: how they re-weight data during the inversion process, and how they incorporate ground truth information. It is much more difficult to compare these aspects of the algorithms, and we defer a discussion of them until a later paper. In the tests shown below, we have attempted to equalize their re-weighting schemes, and we compare only the *relative* event locations each method obtained.

4.4. Comparison of Methods Applied to the Izmit/Duzce Clusters

Engdahl and Bergman (2001) applied HDC to two adjacent event clusters comprising 34 events from the 17the August 1999 Izmit earthquake sequence and 41 events from the 12th November 1999 Duzce sequence. They used Pn and P arrival times from the US Geological Survey National Earthquake Information Center (USGS/NEIC), and ground-truth (local network) locations for a few of the events to constrain the centroid of each cluster. The HDC analysis resulted in approximately 3500 defining phases at 650 stations for the Izmit cluster, and 3200 phases at 600 stations for Duzce. The defining phases covered an epicentral distance range of 3 to 100 degrees in each case.

We have independently applied the other methods to the same arrival time data and GT information. The rules under which these tests were conducted are as follows:

- Each method must be applied to the same set of defining phases used by HDC, using the same phase associations (P vs. Pn).
- Outlier rejection cannot be used.
- Event depths must be fixed to the same depths fixed in the HDC analysis.

Here we show a subset of the results we have obtained thus far and offer a limited interpretation of these results. We show only the Duzce clusters and, with one exception, results which did not incorporate local-network (GT) constraints on the locations. The exception is JHD, which was not run without the GT constraints, but we note that the relative locations are unlikely affected since only two of the Duzce events had GT information. We will focus on relative locations and, to facilitate this, we display each solution registered to the HDC result by shifting it to have the same cluster centroid.

Figure 6 compares two GMEL solutions to the HDC solution. In one case (top panel) the data were weighted equally and the weights were not changed in the inversion. In the second case, GMEL adjusted the data variances in a station-dependent manner by setting the standard error for each station (σ_I) to its r.m.s. residual, as described in the GMEL algorithm above. (However, each σ_I was bounded between 0.5 and 2.0.) It is immediately noticeable from Figure 6 that the GMEL relative locations obtained with this station-dependent weighting are much closer to the HDC locations. In the bottom panel

(station-dependent weighting) the GMEL and HDC relative locations differ more than 2 km for only a few events and never more than 5 km, unlike the top panel. This is because HDC used a similar method of station-dependent weighting of the data and because, apparently, how the data are weighted is an important aspect of multiple-event location in this instance.

Figure 7 compares the DD, PMEL and JHD results to the HDC result for the Duzce cluster. In these other methods, the data weighting is not the same as in HDC. Fixed, uniform weighting was used in PMEL and DD. In DD, however, some possible time differences between events were omitted (when the events were separated by more than 30 km), although the retained pairs were not distance-weighted. The JHD algorithm has a complex weighting scheme, but in this example the weights depend mainly on event-station distance and the effect of this might be roughly comparable to station-dependent weighting.

We see from Figure 7 that the other methods also produce relative locations for the Duzce events much like HDC. However, since the data weighting differs from HDC, the agreement is not as close as the second GMEL case. The JHD results (Figure 7, top) match the HDC ones quite well. The DD epicenters are not quite as close and the PMEL result shows the greatest differences from HDC. A possible reason for this is that PMEL used uniform weighting, which we saw caused significant differences from HDC in the first GMEL case. A second possible reason is that PMEL attempts to fit as much of the data as possible with event location differences. Repeating from above, this would not affect the relative locations in an ideal, noise-free case but the PMEL result suggests that it might affect them in practice. Of particular relevance to this is the fact that this test uses a very large number of stations. For the Duzce cluster, the average number of arrivals per event is high (95) but the number of arrivals per station averages only 5.4. (The situation for Izmit is similar.) This provides the potential for trade-offs between the relative locations and station corrections.

This analysis and its conclusions must be considered preliminary since the comparisons made to date have not considered the errors (confidence regions) in the locations, which might explain some of the larger differences between methods. Since we have limited the comparisons to relative locations, we also have not considered the

agreement of each method with available ground-truth locations, which is ultimately a more important test than the agreement between methods.

5. CONFIDENCE REGIONS AND MODELING ERRORS

In the previous sections we described a new approach to computing event location confidence regions, for the single-event location problem. The confidence region on an event hypocenter is defined as the locus of hypocenters which cannot be rejected based on a hypothesis test that compares the likelihood (L) achieved by any given hypocenter to the maximum likelihood achieved by the best hypocenter. The probability distribution of the test statistic, needed to associate a confidence level with the test, is inferred by Monte-Carlo simulation of the test statistic.

In the multiple-event problem, we seek to obtain a confidence region on an event hypocenter, say $\mathbf{x}_1$, allowing for the fact that the data depend also on the origin time of the event, the location parameters of the other m - 1 events, and the travel-time corrections at the n stations. All these parameters are not uniquely determined by the data, and our uncertainty approach applied to the full multiple-event system will, in principle, automatically account for trade-offs between $\mathbf{x}_1$ and these other parameters. In doing so, the confidence region on $\mathbf{x}_1$ will implicitly account for modeling errors (in this case, uncertainty in the station corrections).

To do this, we must consider two values of the likelihood function, L, given in equation (23). The first, $L_{\max}$, is the value achieved by the maximum-likelihood estimates of the parameters; i.e., L is maximized over all the station and event parameters. The second, which we define as a function of the hypocenter $\mathbf{x}_1$ and denote as $L_{\mathbf{x}_1}$, is the value achieved by maximizing L with respect to all the parameters except the hypocenter of event 1, which is fixed at $\mathbf{x}_1$. The test statistic for defining the confidence region on $\mathbf{x}_1$ is then

$$\tau(\mathbf{x}_1) = \log L_{\max} - \log L_{\mathbf{x}_1} . \tag{27}$$

This statistic essentially compares how well the data can be fit with $\mathbf{x}_1$ fixed vs. free, with all the other parameters free to fit the data. The confidence region on $\mathbf{x}_1$ at confidence level β (e.g. $\beta = 90\%$) is then given by the locus of hypocenters $\mathbf{x}_1$ satisfying

$$F[\tau(\mathbf{x}_1)] \leq \beta \tag{28}$$

$F[\ ]$ denotes the cumulative distribution function (c.d.f.) of a random variable.

The c.d.f. of τ is determined by the assumed distribution of the data errors, e_{ij}. We are currently developing a Monte Carlo scheme for estimating the distribution of τ. The scheme computes $\tau(\mathbf{x}_1)$ from many realizations of synthetic data generated using some assumed true parameters (event locations and station corrections) and varying samples of pseudo-random errors. Applying this approach in the multiple-event case is very computationally intensive since it involves the computation of many multiple-event inversions, one for each point $\mathbf{x}_1$ on a hypocenter grid, using the real data, and two for each realization of synthetic data. We have implemented a basic algorithm and are using it to perform a proof-of-concept of the approach while investigating computational shortcuts and approximations to make it more practical.

5.1. Formulation

We state our formulation of the joint location/calibration inverse problem for the special case of a data set comprising only seismic arrival times, with at most one seismic phase observed for each event-station pair. If there are m seismic events ($i = 1, \ldots, m$) and n seismic stations ($j = 1, \ldots, n$), the inverse problem can be written as given by Equation (20). This equation holds for the station/event pairs for which data have been observed. The unknown parameters of this joint calibration/location inverse problem are the event hypocenters and origin times ($\mathbf{x}_i\, t_i$, $i = 1, \ldots, m$) and the path correction c_{ij}.

The exact nature of the inverse problem depends on how the path corrections c_{ij} are parameterized, and what prior constraints are imposed on them and the event origin parameters. On the latter issue, the problem is purely one of calibration when the event parameters are assumed known and purely one of location when the path corrections are known. In practice, neither set of parameters is completely known or unknown. Two difficult problems in nuclear discrimination, in fact, are the same problem in this

formulation. The first is how to account for uncertainty in a seismic calibration (errors in estimates of c_{ij}) when the calibration events have imperfectly known locations, i.e. ground-truth (GT) levels greater than 0. The second, which is the focus of this project, is how the location error of any particular event is affected by imperfect knowledge of the path corrections. These are two aspects of a complete uncertainty analysis in the joint calibration/location problem.

5.2. Parameterization of Path Corrections

The multiple-event inverse problem is hopelessly ill-posed if the path corrections are not constrained via prior information or a parameterization that reduces the number of independent unknowns, or both. In the basic multiple-event location problem, relevant to small event clusters, the path corrections are assumed to be event-independent, i.e. $c_{ij}=a_j$. The calibration parameters comprise a time term, a_j, for each station. Our work on uncertainty analysis to date has focused on this basic problem. However, our formulation pertains to other ways of parameterizing path corrections, such as with correction functions (or surfaces):

$$c_{ij} = a_j(\mathbf{x}_i) \tag{29}$$

Here, there is an unknown function, a_j (**x**) assigned to each station. It also pertains to the universal parameter functions described by Rodi et al. (2003). When spatial functions are used to parameterize travel-time corrections, however, it is necessary to provide additional prior information on their smoothness, such as with geo-statistical constraints (e.g. Schultz et al., 1998).

5.3. Maximum-Likelihood

As stated previously, our approach to inverse problems and uncertainty analysis is based on likelihood functions. In our location algorithms GSEL and GMEL, we assume that the picking errors are statistically independent and that each has a generalized Gaussian probability distribution of order p as given by Equation (2) (Billings et al., 1994) and modified for σ_{ij} and e_{ij}. For p = 1, the p.d.f. is a Laplace distribution (two-sided

exponential) and for $p = 2$ it is Gaussian. For the small-cluster problem ($c_{ij} = a_j$), this error model implies a likelihood function, L, given by:

$$-\log L = \text{const} + \sum_{ij} \log \sigma_{ij} + \frac{1}{p} \sum_{ij} \frac{1}{(\sigma_{ij})^p} \left| d_{ij} - t_i - T_j(\mathbf{x}_i) - a_j \right|^p . \tag{30}$$

We comment that, with the generalized Gaussian distribution, the negative logarithm of the likelihood function is an L_p norm (to the pth power) of the data residuals. In the Gaussian case it is an L_2 norm, and maximization of the likelihood function with respect to the problem unknowns becomes a least-squares problem.

When correction functions, instead of correction terms, are used to parameterize path corrections, an additional term can be added to L in order to invoke geo-statistical constraints (see Rodi et al., 2003).

5.4. Single-Event Confidence Regions

The uncertainty approach to location error in the single-event location problem has been discussed earlier on this paper. Further details can also be found at Rodi and Toksoz (2001). Our current single-event location algorithm (GSEL) uses a grid-search technique to maximize L with respect to the event hypocenter $\mathbf{x}$.

Flinn (1965), Evernden (1969) and Jordan and Sverdrup (1981) developed the methodology for hypocentral confidence regions for the case of Gaussian data errors ($p = 2$) and using a linear approximation to the travel-time functions, T_j. Their methods can be formulated in terms of hypothesis testing using a likelihood ratio as the test statistic. Doing so allows us to define confidence regions under less restrictive assumptions.

Confidence regions can be found on all four location parameters simultaneously (including origin time) or on any subset of parameters. For a 3-D, hypocentral confidence region on $\mathbf{x}$, the relevant test statistic is given by Equation 11. This statistic compares the value of L achieved by maximizing it over all unknown parameters to the value of L when maximized over only t and σ, with the hypocenter fixed to any given $\mathbf{x}$. Given a confidence level β, we can reject $\mathbf{x}$ at that level if $\tau(\mathbf{x},\mathbf{d})$ is greater than some critical value, τ_β. This critical value is determined by the probability distribution of $\tau(\mathbf{x},\mathbf{d})$, as

induced by the errors in **d**. If **x** is *not* rejected, it is *inside* the confidence region for the specified confidence level. That is, the confidence region comprises the hypocenters **x** satisfying:

$$\tau(\mathbf{x},\mathbf{d}) \leq \tau_\beta \tag{31}$$

Under Flinn's assumptions, τ is F-distributed and τ_β. can be found directly from the F cumulative distribution function (c.d.f.). The locus of points **x** satisfying equation (31) fill an ellipsoid whose axis lengths and orientations are easily found.

With our more general assumptions, τ does not necessarily have a well-known probability distribution and the geometry of the confidence region cannot be found with analytic formulas. Therefore, we use numerical techniques to find τ_β. and the values of **x** that satisfy equation (31). Our confidence region method becomes a two-step procedure:

1. Sample $\tau(\mathbf{x},\mathbf{d})$ at points **x** on a dense 3-D hypocenter grid. This entails maximizing L with respect to origin time and σ for each point on the grid, and comparing each such value to the maximum L over all parameters (i.e. the maximum likelihood solution).

2. Perform a Monte Carlo simulation to find τ_β. for various β. This entails computing $\tau(\mathbf{x}^{tru},\mathbf{d}^{syn})$ for an assumed "true" hypocenter $\mathbf{x}^{tru}$ and many realizations of synthetic data, $\mathbf{d}^{syn}$. Each realization $\mathbf{d}^{syn}$ is constructed by adding a realization of pseudo-random errors to arrival times calculated for $\mathbf{x}^{tru}$. The random noise is generated according to the assumed error distribution using an assumed "true" value of σ.

Both steps of this procedure can be computationally intensive. The first step (likelihood sampling) requires maximizing L for all hypocenters on a 3-D grid, but only for the real data. The second step (Monte Carlo simulation) entails maximizing L twice, with **x** fixed and free, for each of many realizations of synthetic data. We have found that about 200–300 realizations of synthetic data yield stable results.

Confidence regions on other event parameters are obtained in an analogous manner, and share some of the calculations needed for hypocentral regions.

Examples

An example of our approach, which demonstrates the importance of accounting for travel-time nonlinearity, is shown in Figure 8. This figure displays confidence regions for an earthquake in Honshu, Japan, as determined from mostly first-arrival, teleseismic data obtained by the International Monitoring System (IMS) seismic network. Each figure shows the 2-D confidence region of the event epicenter (at varying confidence levels) and cross-sections through the 3-D confidence region on the event hypocenter. The IASP91 travel-time tables were used in these computations, and the best-fitting location of the event is in the mantle of the IASP91 model (depth = 44 km). The epicenter confidence region for each confidence level (right of figure) is nearly elliptical in shape. The corresponding hypocentral confidence region, however, is quite non-ellipsoidal and extends upward into the crust (left and middle frames). This is an effect of nonlinearity that is not accounted for with the conventional Gaussian/linear formulas. The cause of it is the velocity difference, and consequent difference in the sensitivity of travel-time to event depth (dT/dz), across the Moho. The conventional confidence ellipsoid formulas in this case would assume the mantle dT/dz holds everywhere, precluding the asymmetry seen in Figure 8. This example illustrates only one simple nonlinear phenomenon that can occur even with well-recorded events, 1-D Earth models, and Gaussian data errors (our computations used a Gaussian error model).

A second example of Monte Carlo confidence regions is shown in Figure 9. These are for an event (no. 91135030) from the 1991 Racha, Georgia earthquake sequence, studied by Myers and Schultz (2000). The data set comprised P wave arrival times from 6 stations, obtained from ISC and supplemented with picks made by F. Ryall of LLNL. The stations cover an epicentral distance range of 6° to 22°.

We computed confidence regions using two different assumptions about the probability distribution of picking errors. The top frames of Figure 9 assumed a Gaussian distribution ($p = 2$) while the bottom frames assumed a Laplace distribution ($p = 1$). Even in the Gaussian case we see that the confidence regions are blatantly non-elliptical, especially the epicentral confidence regions (rightmost frames). As with the Honshu events, a big contributor to non-ellipticity is the nonlinearity of the travel-times with respect to event depth. This factor is even more significant in this example because the

event depth is so poorly constrained by the sparse network available. We also see that Laplace-error confidence regions (bottom frames of Figure 9) are blockier and larger than Gaussian-error ones, which can be attributed to the differing properties of L_1 and L_2 error norms.

We note that the confidence regions in Figure 9 do not cover the GT location inferred by Myers and Schultz (the circles in each frame), and our maximum-likelihood location is biased by about 30 km. This is because we did not apply travel-time corrections to the data, or attempt to account for them as modeling errors in the computation of the confidence regions.

5.5. Multiple-Event Location Confidence Regions

Our multiple-event location algorithm (GMEL) solves jointly for the location parameters, x_i and t_i, of m events, and the travel-time correction terms, a_j, for n stations. It finds maximum-likelihood solutions for these unknown parameters, using the multiple-event, multiple-station version of the likelihood function. Like its single-event predecessor, GMEL implements the generalized Gaussian error model, this time allowing unknown station-specific scale factors, σ, on the standard errors (Equation 22).

Hard bounds on all the unknown parameters, including upper and lower bounds on the a_j, are allowed. GMEL finds the maximum-likelihood solutions for the parameters by iterating over alternating loops over events, to update event locations with station terms fixed, and over stations, to update the station terms and station scale factors, with the event locations fixed. Grid-search is used to find the optimal event locations (see Rodi et al., 2002).

To define confidence regions on one of the m events in the multiple-event location problem, we have adapted our location uncertainty approach as follows. Let us denote the likelihood function for the multiple-event location problem $L(\mathbf{x}_1, t_1, \mathbf{x}_2, t_2, \ldots, \mathbf{x}_m, t_m, a_1, \sigma_1, a_2, \sigma_2 \ldots, a_n, \sigma_n; \mathbf{d})$, where the vector $\mathbf{d}$ now contains the data for all events and stations. Let us arbitrarily identify the event of interest to be the first (i =1). To define a confidence region on its hypocenter, $\mathbf{x}_1$, our uncertainty paradigm tells us to define a test-statistic, $\tau(\mathbf{x}_1, \mathbf{d})$ that compares two values of the likelihood function. The

first is the value achieved by maximizing L over all event and station parameters: $\mathbf{x}_1$, t_1, $\mathbf{x}_2$, t_2, …, $\mathbf{x}_m$, t_m and a_1, σ_1, a_2, σ_2, …, a_n, σ_n. The second value of L is that achieved by maximizing L over all the parameters except $\mathbf{x}_1$, which is held fixed. Maximization of L in both cases is performed subject to available GT constraints on any of the event locations. The formula for $\pi(\mathbf{x}_1, \mathbf{d})$ is analogous to equation (11).

Given this, our method for computing a confidence region on $\mathbf{x}_1$ is a two-step process analogous to the single-event case:

1. Sample $\pi(\mathbf{x}_1, \mathbf{d})$ or values of $\mathbf{x}_1$ on a dense grid.
2. Compute Monte Carlo realizations of $\tau(\mathbf{x}_1^{\text{tru}}, \mathbf{d}^{\text{syn}})$ for many random realizations of (multiple-event) synthetic data, $\mathbf{d}^{\text{syn}}$.

Each computation of τ, whether for real or synthetic data, now entails computing the solution to a multiple-event inverse problem involving the data and parameters for all events and stations.

This approach to location confidence regions implicitly accounts for the uncertainty in $\mathbf{x}_1$ induced by its trade-off with the travel-time corrections, a_j, as they are constrained by the multiple-event data set and GT information on the other events. It avoids the need for an explicit uncertainty model for the station terms, which would then be treated as modeling errors added to the picking errors. This implicit approach takes into account correlations and the finite accuracy of GT events, and handles nonlinearity and non-Gaussian data errors.

Example from the Izmit (Turkey) earthquake sequence

We show some preliminary results of our multiple-event uncertainty approach applied to the 17 August 1999 Izmit, Turkey, earthquake and 32 of its aftershocks. The data set contains approximately 3500 Pn and teleseismic P arrival times from 640 stations from National Earthquake Information Center (NEIC), which we obtained courtesy of R. Engdahl and the IASPEI Working Group on Multiple-Event Location. Previously, Engdahl and Bergman (2001) applied the hypocentroidal decomposition (HDC) method to this cluster and a similar cluster from the 12 November 1999 Duzce earthquake sequence. We show confidence regions that we have computed for one of the events in the Izmit cluster.

The event we consider is a well-recorded aftershock occurring on 31 August 1999, and has 138 defining phases in the data set (compared to 388 defining phases for the Izmit mainshock). The top/left frame of Figure 10 shows the single-event, epicentral confidence region for this event, computed independently of the other events and with zero travel-time corrections at the stations ($a_j = 0$). Gaussian data errors were assumed and the depth of the event was fixed. The circle marks a GT5 location for the aftershock determined from a local seismic network (provided by R. Engdahl for the IASPEI Working Group). We see that the single-event confidence region is displaced by about 20 km from the GT location, and even the 98% confidence region does not include the GT location. This is expected since travel-time corrections were not applied to the AK135 travel-time tables that were used in the computations.

The other frames of Figure 10 show different versions of multiple-event confidence regions for the same Izmit aftershock. In each case, the event was located jointly with the 32 other events and the estimation of travel-time corrections at the roughly 640 stations in the data set. These other 32 events, which include the Izmit mainshock, play the role of calibration events. The different versions of multiple-event confidence regions have to do with different GT constraints we used for the calibration events. The top/right frame assumed that the Izmit mainshock is a GT0 event. Thus, the mainshock was constrained to its local network solution while the other event epicenters were free to vary. The bottom/left frame also assumed the Izmit mainshock was the only GT event, but this time it was assumed to be GT5, i.e. its epicenter was constrained to be within 5 km of the local network solution. The bottom/right frame uses no GT constraints on any of the events. In each frame, the maximum-likelihood solution for the aftershock found by GMEL is in the center of the plot (at zero northing and easting) although it is not marked. In the top/right frame this is near the center of the confidence region, but this is not the case for the bottom two frames.

Comparing the top frames of Figure 10, we see that the multiple-event confidence region for the aftershock is much closer to its local network solution than was the case for the single-event confidence region. This is because the requirement that the event locations of all the events be self-consistent with the station travel-time corrections, combined with the GT0 constraint on the mainshock, removes much of the bias in the

GMEL location that we saw in the single-event case. However, the local network solution (which might be in error by as much as 5 km) is a little outside the confidence regions at each confidence level. In the bottom/left frame, which assumed the mainshock was GT5, the confidence region for the aftershock is slightly larger, as expected, and comes closer to including the local network solution. In the bottom/right frame, we see that the multiple-event confidence region resulting from no GT constraints is much larger. This largely reflects the fact that, in small-aperture multiple-event location, there is a strong trade-off between the centroid of the cluster and the station corrections; i.e. absolute events locations cannot be determined without GT constraints. In the linear theory (Jordan and Sverdrup, 1981) this indeterminacy is total, and yet the confidence region in the bottom/right frame is finite. The probable explanation for this is that, for computational reasons, we limited the extent to which event locations could trade off with station travel-time corrections to 15 km. However, another reason might be that we imposed hard bounds on the station travel-time corrections, ±10 sec relative to AK135 times, which is a reasonable assumption. It may be that this constraint also has limited the size of the confidence regions in the bottom/right frame, and may also explain why they are not centered on the ones computed *with* a GT constraint on the mainshock.

These results, while preliminary, seem to support the validity of our implicit approach to modeling errors. Further, the results suggest the possibility that, even in this simplest of joint location/calibration problems, prior constraints on both reference event locations and travel-time discrepancies from global 1-D models might be an important element in location error assessment.

6. CONCLUSIONS AND RECOMMENDATIONS

We have developed a general theoretical and computational framework for characterizing the uncertainty in seismic event locations in the context of the joint inverse problem of multiple-event location and travel-time calibration. In both single and multiple event location algorithms (GSEL and GMEL) that we have developed, our approach accommodates Gaussian and non-Gaussian assumptions about the picking errors in seismic arrival times, and avoids linearization of the forward travel-time problem. Our comparison of GMEL to other multiple-event location methods (HDC, DD, JHD and PMEL) has helped to understand the similarities and differences among these methods, in theory and in practice. This is an important step in the development of GMEL and its new approach to location uncertainty. The most important new element of our approach is that it takes implicit account of the errors in travel-time forward models. This avoids a difficult task when the calibration and location problems are addressed separately, i.e. determining realistic uncertainty estimates in seismic calibration parameters, and then incorporating them into a single-event event location algorithm. However, our approach introduces a new difficulty in the fact that it is computationally intensive. We are investigating computational short-cuts that will make our method more practical. Further, we are looking at how best to implement the approach with more general parameterizations of path corrections, including travel-time correction surfaces, so that the method is not restricted to small event clusters.

7. REFERENCES

Aktar, M., M. Ergin, S. Özalaybey, C. Tapirdamaz, A. Yörük and F. Bicmen, 2000. A lower-crustal event in the northeastern Mediterranean: The 1998 Adana earthquake (M_w = 6.2) and its aftershocks, *Geophys. Res. Letters*, **27**, 2361--2364.

Billings, S.D., M.S. Sambridge and B.L.N. Kennett, 1994. Errors in hypocenter location: picking, model and magnitude dependence, *Bull. Seism. Soc. Am.*, **84**, 1978–1990

Dewey, J.W., 1971. Seismicity studies with the method of joint hypocenter determination, Ph.D. Thesis, University of California, Berkeley.

Douglas, A., 1967. Joint epicentre determination, *Nature*, **215**, 47-48.

Engdahl, E.R. and E.A. Bergman, 2000. Identification and validation of reference events within the area being regionally monitored by IMS stations in Asia and North Africa, *Proceedings*, 22nd Annual DoD/DOE Seismic Research Symposium, New Orleans.

Engdahl, E.R. and E.A. Bergman, 2001. Validation and generation of reference events by cluster analysis, *Proceedings*, 23rd Annual DoD/DOE Seismic Research Review, Jackson Hole, Wyoming.

Evernden, J.F., 1969. Precision of epicenters obtained by small numbers of world-wide stations, *Bull. Seism. Soc. Am.*, **59**, 1365–1398.

Flinn, E.A., 1965. Confidence regions and error determinations for seismic event location, *Rev. Geophys.*, **3**, 157–185.

Jeffreys, H., 1932. An alternative to the rejection of observations, *Mon. Not. R. Astr. Soc, Geophys. Suppl.*, **2**, 78--87.

Jordan, T.H. and K.A. Sverdrup, 1981. Teleseismic location techniques and their application to earthquake clusters in the south-central Pacific, *Bull. Seism. Soc. Am.*, **71**, 1105-1130.

Lilwall, R.C. and A. Douglas, 1970. Estimation of P-wave travel-times using the joint epicentre method, *Geophys. J. R. Astr. Soc.*, **19**, 165-181.

Myers, S.C., and C.A. Schultz, 2000. Improving sparse network seismic locations with Bayesian kriging and teleseismically constrained calibration events, *Bull. Seism. Soc. Am.*, **90**, 199–211.

Pavlis, G.L and J.R. Booker, 1980. The mixed discrete continuous inverse problem: application to the simultaneous determination of earthquake hypocenters and velocity structure, *J. Geophys. Res.* **85**, 4801-4810.

Pavlis, G.L, and J.R. Booker, 1983. Progressive multiple event location (PMEL), *Bull. Seism. Soc. Am.* **73**, 1753-1777.

Rodi, W., E.R. Engdahl, E.A. Bergman, F. Waldhauser, G.L. Pavlis, H. Israelsson, J.W. Dewey and M.N. Toksöz, 2002. A new grid-search multiple-event location algorithm and a comparison of methods, *Proceedings*, 24th Annual Seismic Research Review, Ponte Vedra Beach, Florida, 403–411.

Rodi, W., S.C. Myers and C.A. Schultz, 2003. Grid-search location methods for ground-truth collection from local and regional seismic networks, *Proceedings*, 25th Annual Seismic Research Review, Tucson, Arizona (this volume).

Rodi, W. and M.N. Toksöz, 2000. Grid-search techniques for seismic event location, *Proceedings*, 22nd Annual DoD/DOE Seismic Research Symposium, New Orleans.

Rodi, W. and M.N. Toksöz, 2001. Uncertainty analysis in seismic event location, *Proceedings*, 23rd Annual DoD/DOE Seismic Research Review, Jackson Hole, Wyoming.

Sambridge, M., 1999. Geophysical inversion with a neighbourhood algorithm—1. Searching a parameter space, *Geophys. J. Int.*, **138**, 479–494.

Schultz, C.A., S.C. Myers, J. Hipp and C.J. Young, 1998. Nonstationary Bayesian kriging: a predictive technique to generate spatial corrections for seismic detection, location, and identification, *Bull. Seism. Soc. Am.*, **88**, 1275–1288.

Waldhauser, F. and W.L. Ellsworth, 2000. A double-difference earthquake location algorithm: method and application to the northern Hayward fault, California, *Bull. Seism. Soc. Am.* **90**, 1353-1368.

Wolfe, C.J., 2002. On the mathematics of using difference operators to relocate earthquakes, *Bull. Seism. Soc. Am.*, **92**, 2879 – 2892.

FIGURES

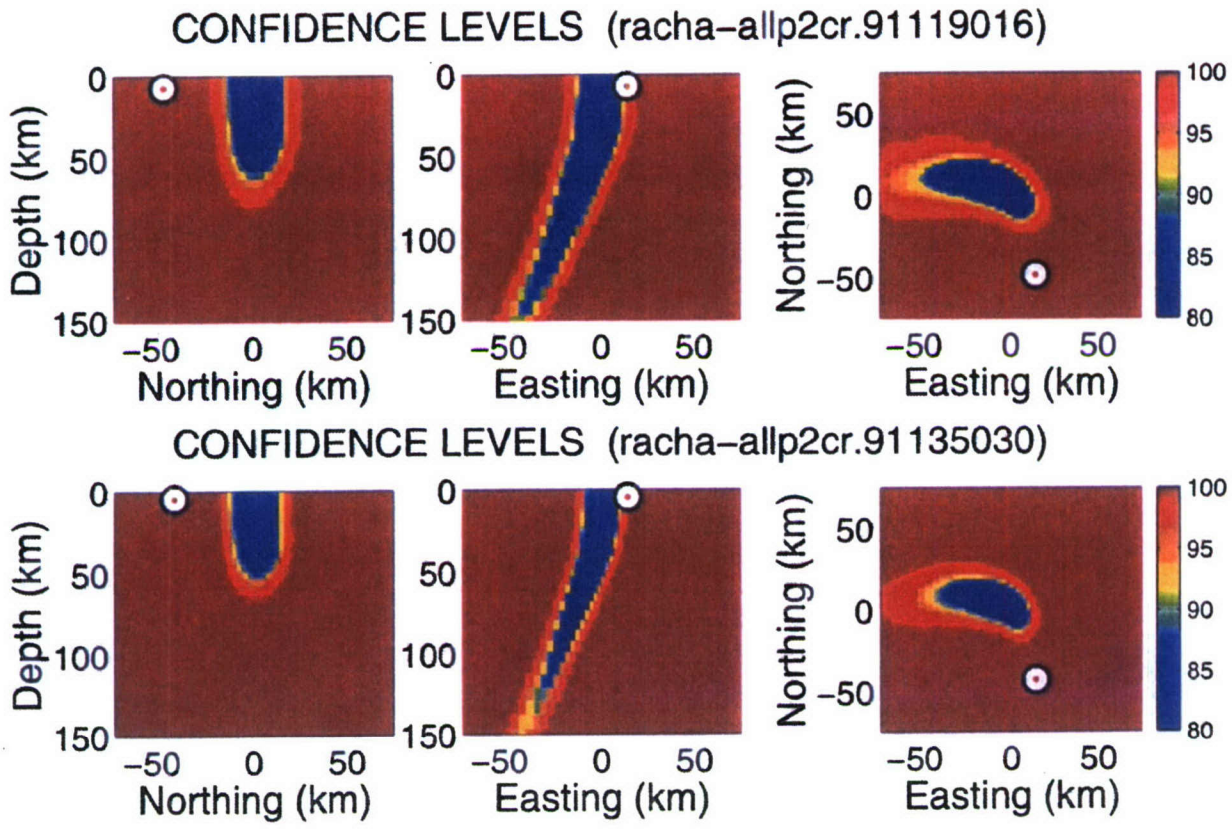

Figure 1: Confidence level vs. location for two events from the 1991 Racha earthquake sequence, obtained from 5 and 6 regional P wave arrival times, respectively. Picking errors are assumed to be Gaussian (p = 2). Left and center: Cross-sections of confidence level vs. hypocenter taken through the maximum-likelihood location. Right: Confidence level vs. epicenter. Each contour of constant confidence level is the boundary of the 3-D hypocenter (left and center) or epicenter (right) confidence region at that level. The best-fitting (ML) epicenter for each event is zero northing and easting. Circles mark the local network (ground-truth) locations for the events.

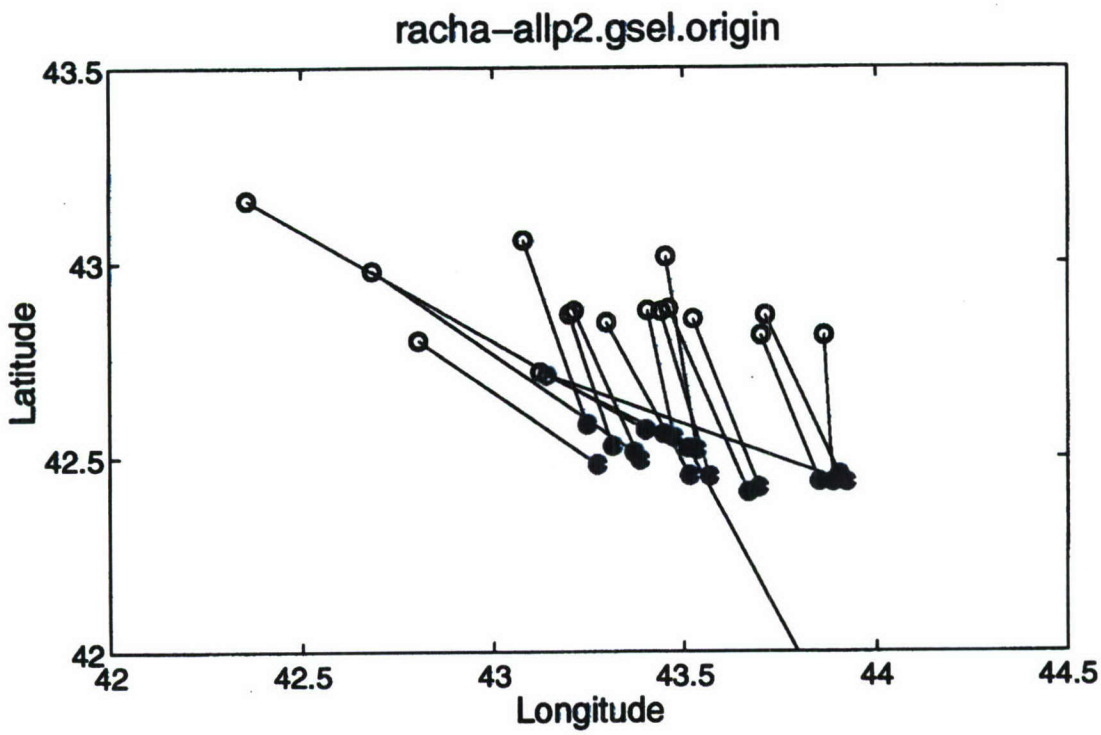

Figure 2: Epicenters of 18 earthquakes from the 1991 Racha sequence. The open circles are the locations determined from P wave arrival times observed at a sparse regional network. The filled circles are ground-truth locations.

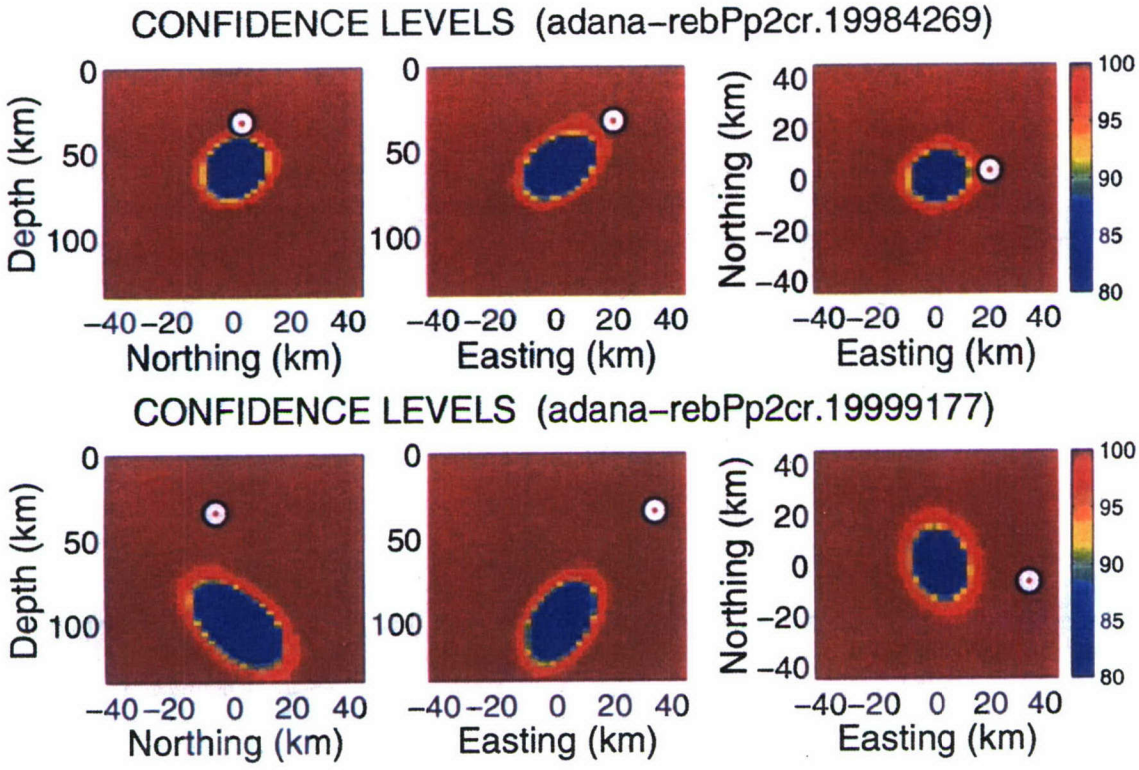

Figure 3: Confidence level vs. location for the 27 June 1998 Adana earthquake and a 4 July 1998 aftershock. A Guassian error model was assumed. The circles mark the local network (ground-truth) locations for the events.

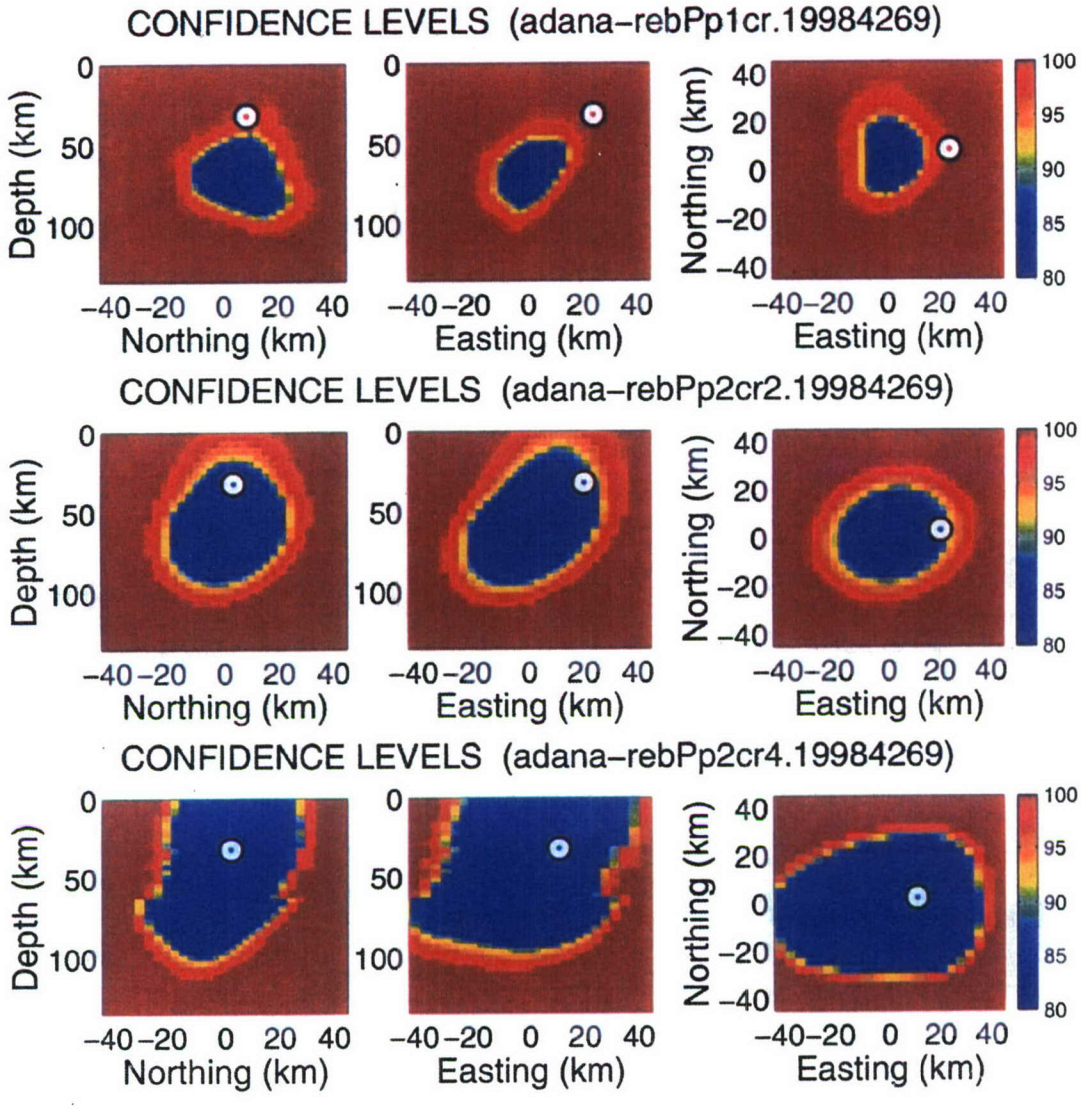

Figure 4: Confidence levels for the 27 June 1998 Adana mainshock derived from alternative error models. Top: exponential distribution ($p = 1$, $\sigma = 1.0$ sec). Center: Gaussian distribution with inflated variance ($p = 2$, $\sigma = 2.0$ sec). Bottom: Gaussian error model with unknown station corrections ($p = 2$, $\sigma = 1.0$. $-2 < c_i < +2$ sec).

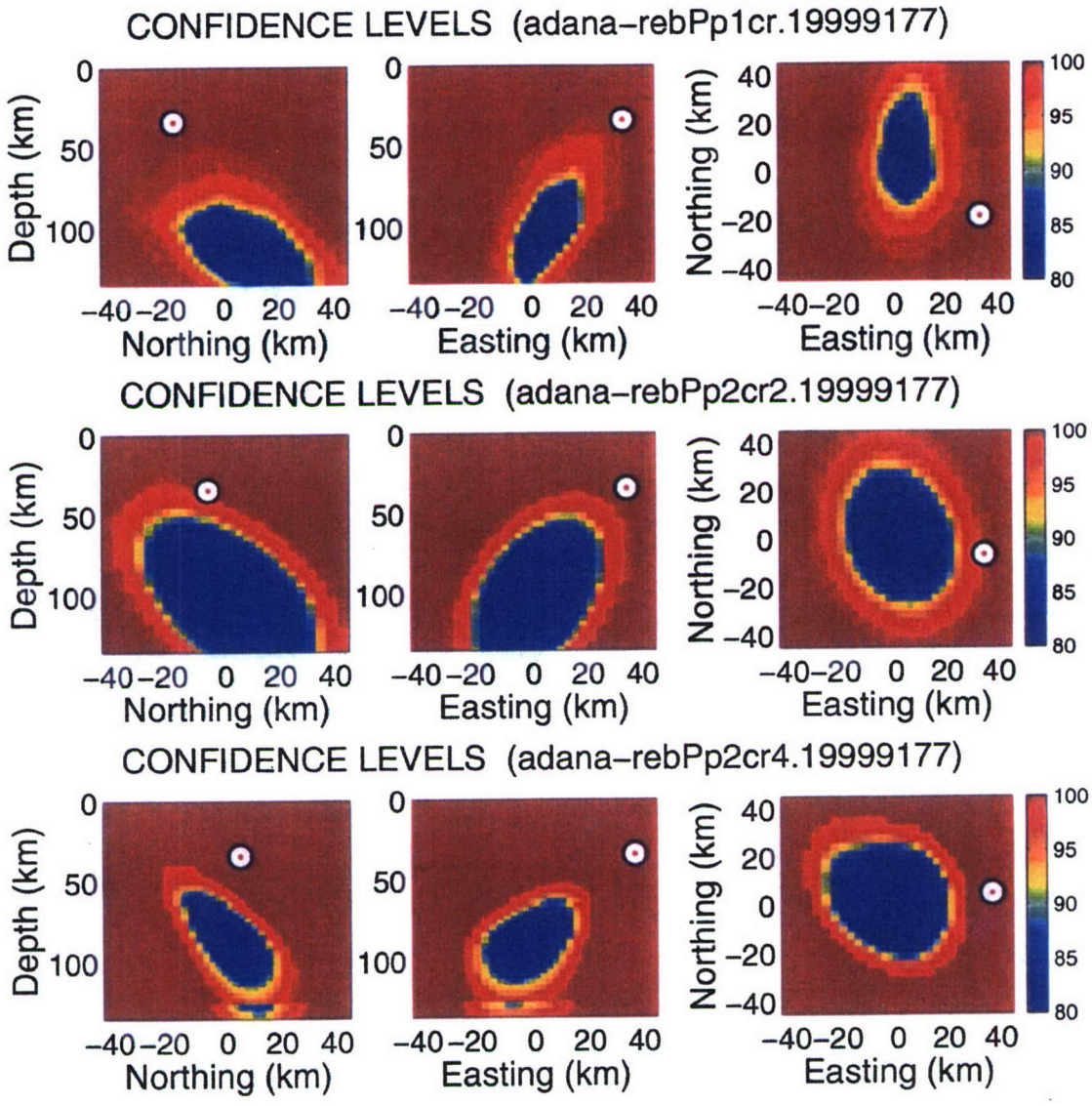

Figure 5: Same as Figure 4 but for the 4 July 1998 Adana aftershock.

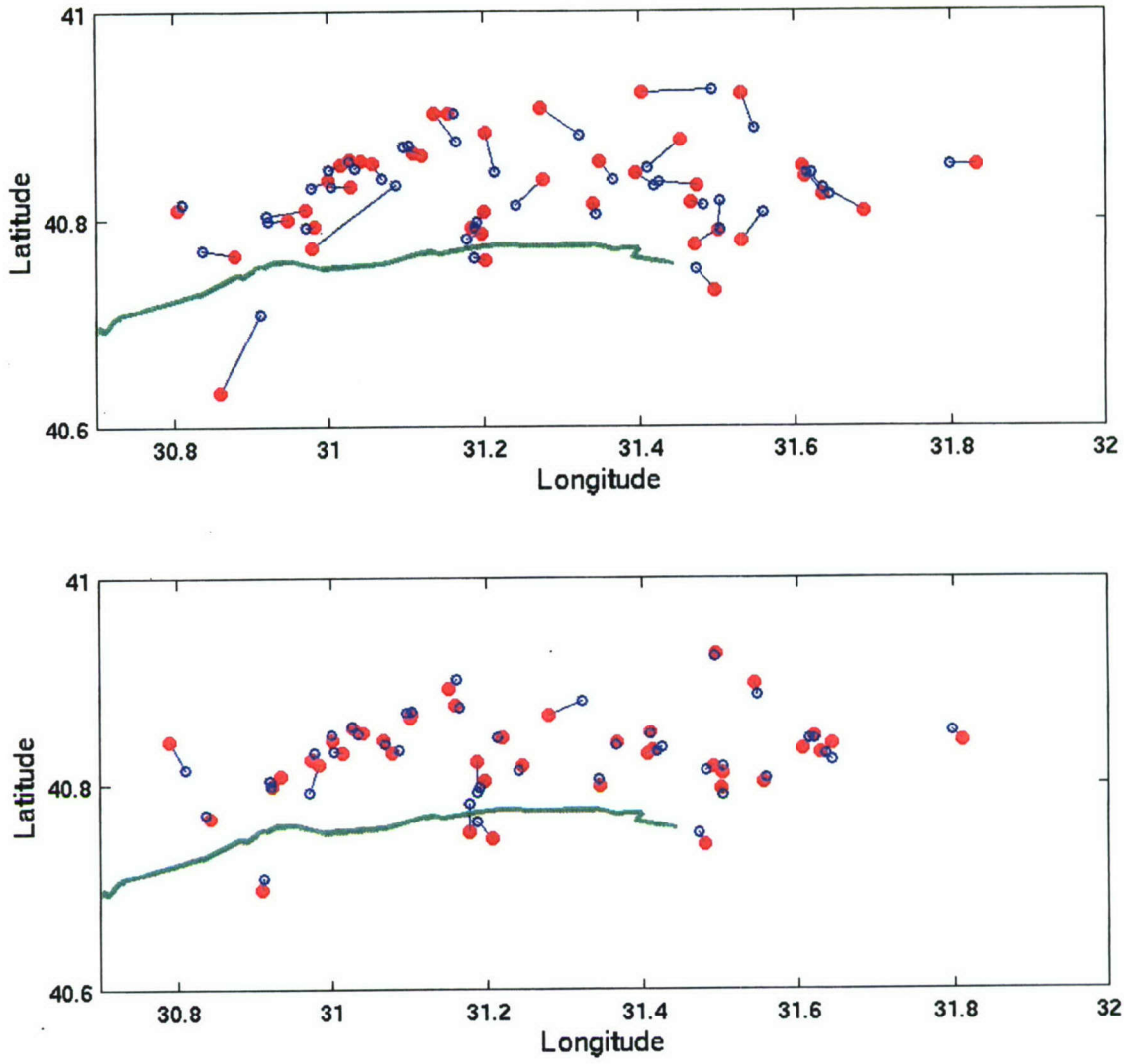

Figure 6: GMEL relative locations (filled circles) vs. HDC relative locations (unfilled) for the Duzce cluster. Two GMEL results are shown. *Top*: using fixed, uniform weighting of the data. *Bottom*: allowing a station-dependent weighting derived from the r.m.s. residual at each station. A line connects the HDC location of each event to the GMEL location of the same event. The surface trace of the North Anatolian Fault is shown for reference.

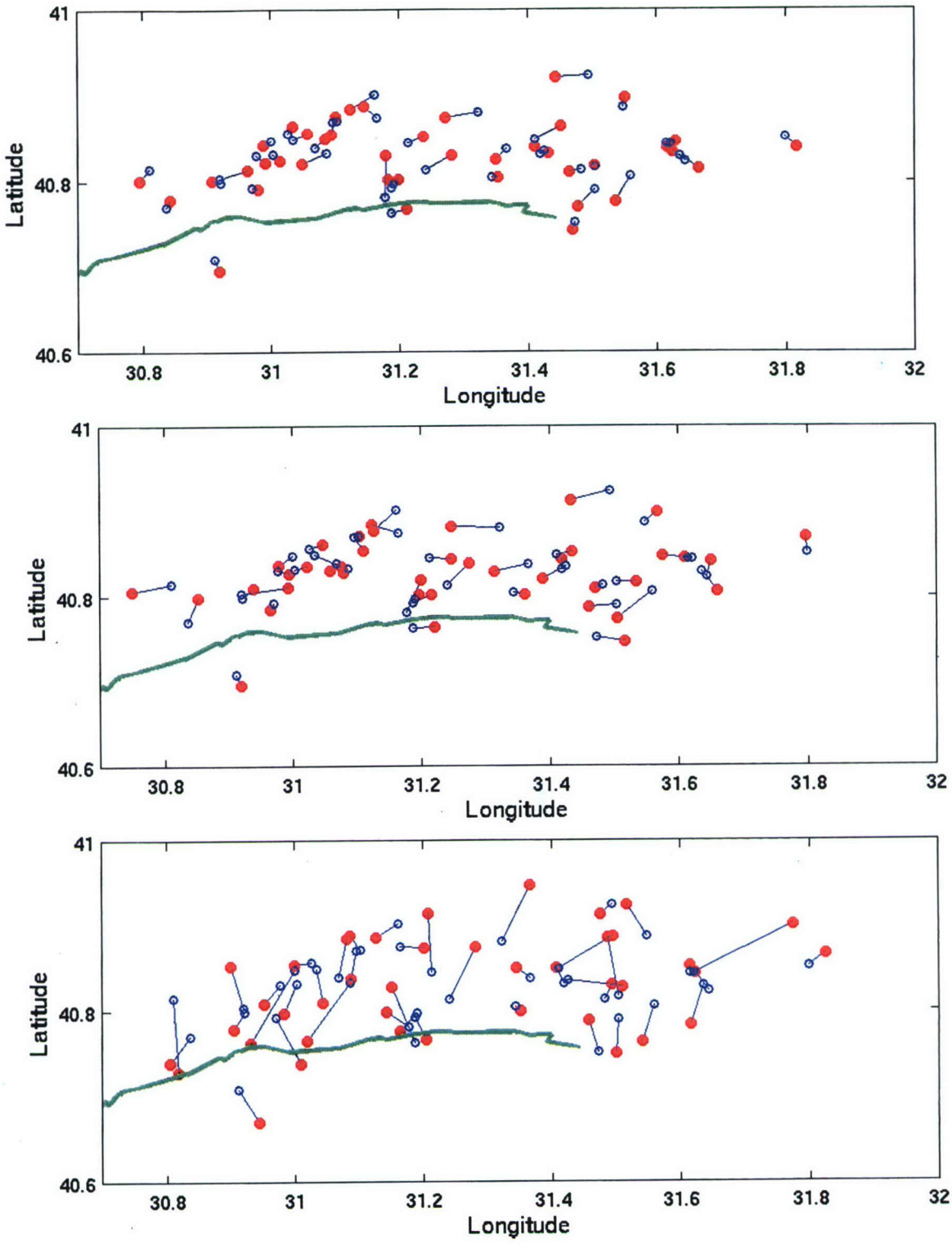

Figure 7: JHD, DD, and PMEL relative locations (filled circles) vs. HDC relative locations (unfilled) for the Duzce cluster. *Top*: JHD results, run with variable data weights. *Middle*: DD results, run with fixed, uniform data weights but with some differences omitted (see text). *Bottom*: PMEL results, run with fixed, uniform data weights.

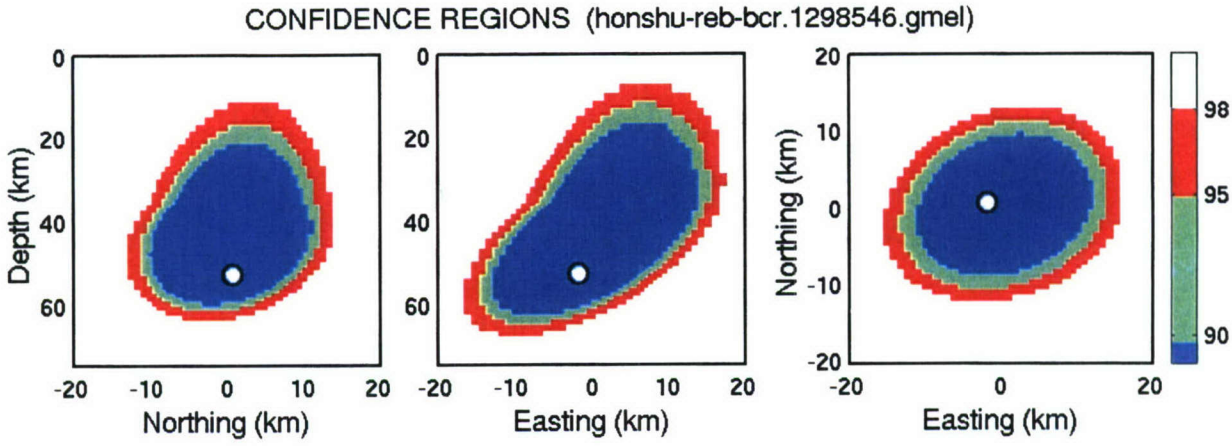

Figure 8: Confidence regions for a 1998 earthquake in Honshu determined using 23 REB arrival time data (2 Pn, 1 Sn, 18 P, 1 S and 1 PKP). Confidence regions at three confidence levels are displayed: 90% (inner (blue) shaded area), 95% (combined inner and central (blue and green) shaded areas), and 98% (inner, central and outer (blue, green and red) shaded areas). *Left and middle frames*: cross-sections through 3-D confidence regions on the event hypocenter. *Right frame*: 2-D confidence regions on the event epicenter. The circle marks the event location reported in the REB.

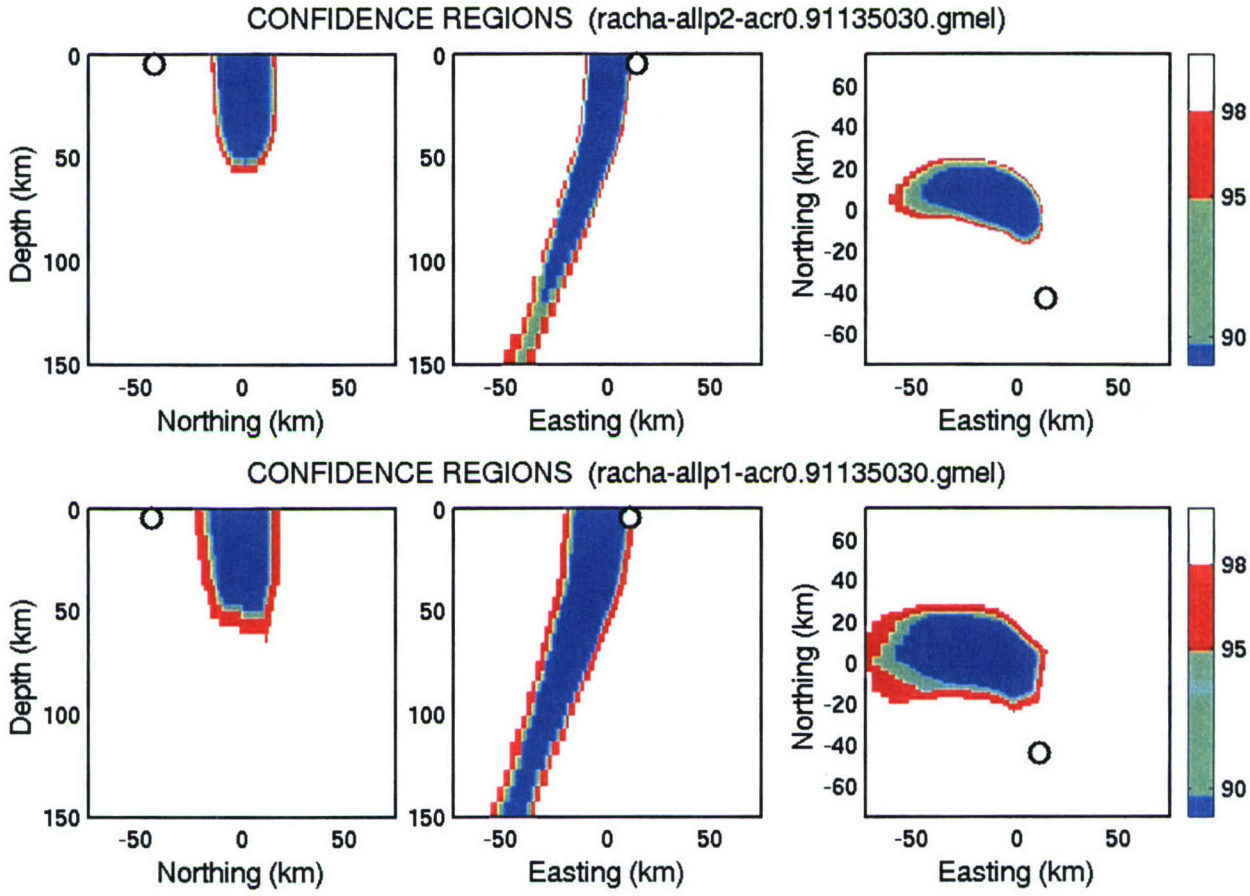

Figure 9: Confidence regions (at 90, 95 and 98%) for an event from the 1991 Racha earthquake sequence. The data set used comprised first-arrival P wave times at 6 regional stations, with epicentral distances varying from 6 to 22°. The top frames are based on the assumption of Gaussian picking errors ($p = 2$). The bottom frames assume the errors are from a Laplace (two-sided exponential) ($p = 1$). *Left and middle*: Cross-sections of 3-D confidence regions on the event hypocenter, taken through the maximum-likelihood location. *Right*: Confidence regions on epicenter. The circle in each plot marks the local network location for the event reported by Myers and Schultz (2000). The maximum-likelihood epicenter is at zero northing and zero easting.

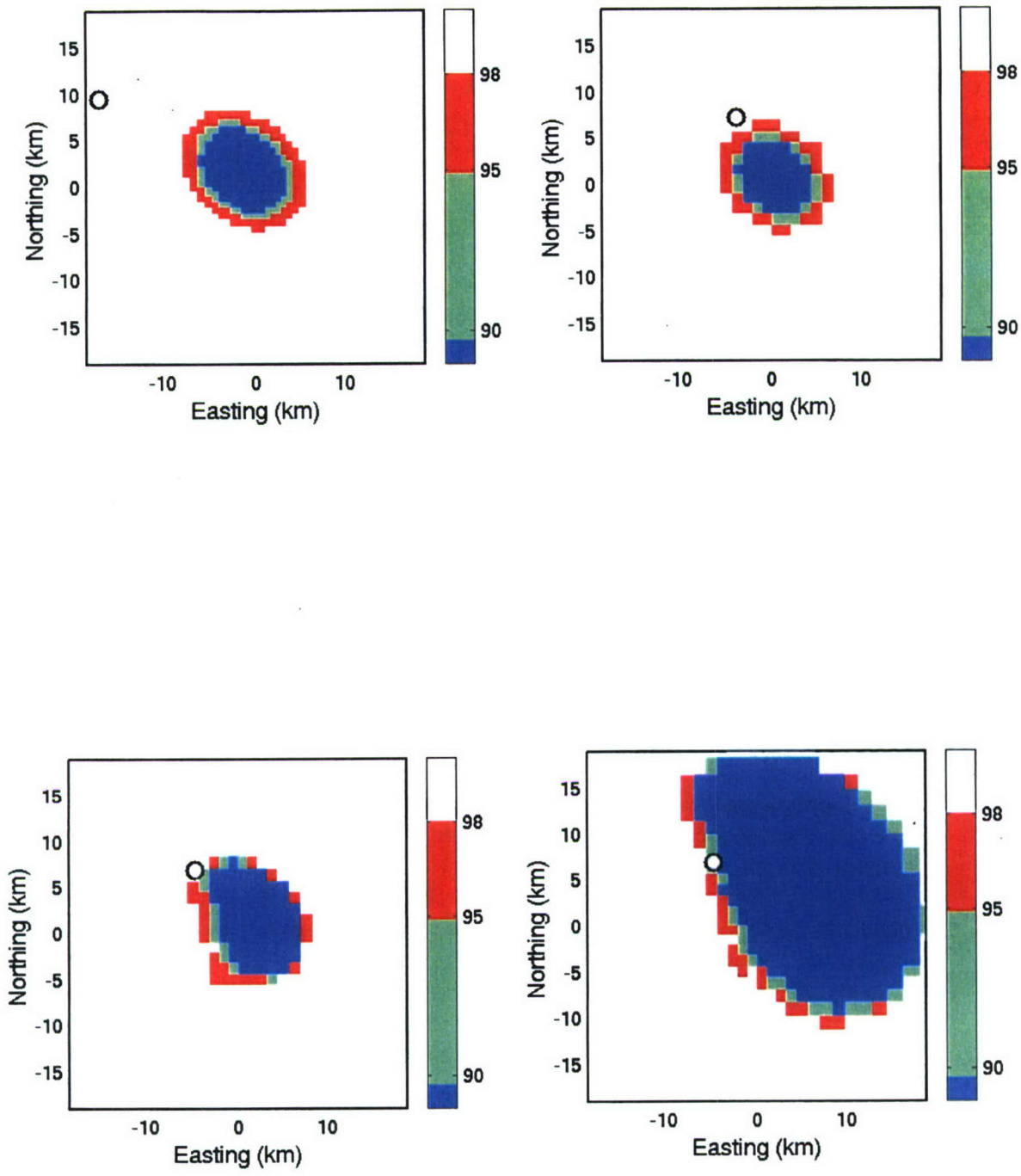

Figure 10: *Top left*: Single-event epicentral confidence regions (at 90, 95 and 98% confidence) for a well-recorded aftershock of the 1999 Izmit, Turkey earthquake. Station travel-time corrections were assumed to be zero. *Top right and bottom*: Multiple-event confidence regions for the same event. *Top right*: computed with the Izmit mainshock constrained as a GT0 event. *Bottom left*: computed with the Izmit mainshock constrained as a GT5 event. *Bottom right*: computed using no GT constraints on any events. In each frame, the circle marks the ground-truth (GT5) epicenter for the aftershock.

DISTRIBUTION LIST
DTRA-TR-06-34

DEPARTMENT OF DEFENSE

DEFENSE TECHNICAL INFORMATION CENTER
8725 JOHN J. KINGMAN ROAD,
SUITE 0944
FT. BELVOIR, VA 22060-0944
2 COPIES ATTN: DTIC/OCA

DEFENSE THREAT REDUCTION AGENCY
8725 JOHN J. KINGMAN ROAD
MS 6201
FT. BELVOIR, VA 22060-6218
2 COPIES ATTN: CW/S. MANGINO

DEPARTMENT OF DEFENSE CONTRACTORS

ITT INDUSTRIES
ITT SYSTEMS CORPORATION
1680 TEXAS STREET, SE
KIRTLAND AFB, NM 87117-5669
2 COPIES ATTN: DTRIAC
ATTN: DARE

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
DEPTARTMENT OF EARTH, ATMOSPHERIC &
PLANETARY SCIENCES
EARTH RESOURCES LABORATORY
CAMBRIDGE, MA 02139
ATTN: N. TOKSOZ